



FACULTAD DE CIENCIAS

Reconstrucción del momento transverso
de un mediador de Materia Oscura
utilizando una red neuronal artificial
*(Reconstruction of the transverse
momentum of a Dark Matter mediator
using an artificial neural network)*

Trabajo de Fin de Grado para acceder al
GRADO EN FÍSICA

Autor: Rubén López Ruiz
Director: Pablo Martínez Ruiz del Árbol

Junio - 2021

Agradecimientos

Me gustaría agradecer en primer lugar a mi director de TFG, Pablo, toda la ayuda, apoyo y conocimiento prestado en la realización de este trabajo. Gracias también por ofrecerme una primera oportunidad e introducirme en el campo de la Física de Altas Energías.

En segundo lugar, también aprovecho para agradecer a mis padres y a mi novia su apoyo en los momentos más difíciles de esta carrera. Por último, gracias a mis amigos por estar siempre ahí y ser un pilar dentro y fuera de la Universidad.

Resumen

La naturaleza de la materia oscura presente en el Universo sigue siendo un misterio a día de hoy. El experimento CMS mantiene un intenso programa de búsqueda de partículas candidatas a conformar la Materia Oscura en las colisiones protón-protón del LHC. Dichas partículas interactúan débilmente con el detector siendo únicamente visibles a través del momento faltante en la colisión. Este trabajo está enmarcado en esta búsqueda de materia oscura. En concreto se estudia el uso de las redes neuronales artificiales en modo regresión para predecir el momento transversal de la partícula de materia oscura en cada suceso. Este estudio se realiza para dos procesos de producción de materia oscura: en asociación con un par quark top - quark antitop y solamente con un quark top decayendo leptónicamente. Además, se analiza la calidad del método mediante una comparación con otros métodos analíticos desarrollados previamente.

Abstract

The nature of Dark Matter present in our Universe is still a mystery. CMS is an experiment with an intensive programme regarding the search of Dark Matter candidate particles in proton-proton collisions at LHC. These particles barely interact with detectors, and can only be seen through the missing energy of the collision. This project is directly related to this Dark Matter search. It specifically analyzes the use of regressive artificial neural networks in order to predict the transverse momentum of a Dark Matter particle in each collision. This study is performed for two Dark Matter production processes: in association with a pair top quark - antitop quark and with only a top quark that decays to leptons. To conclude this project the quality of the method is analyzed by performing a comparison with other previous analitic methods.

Índice general

Resumen / <i>Abstract</i>	III
1. Introducción	1
1.1. Historia y evidencias de Materia Oscura	1
1.2. El Modelo Estándar de la Física de Partículas	2
1.3. La Materia Oscura en Física de Partículas.	4
1.4. Experimento CMS	5
1.5. Búsquedas de Materia Oscura	6
2. Procesos de producción de Materia Oscura con quarks top	9
2.1. Simulación de sucesos	9
2.2. Interacción top - antitop estándar ($t\bar{t}$)	9
2.2.1. Reconstrucción del momento de los neutrinos en el proceso $t\bar{t}$ estándar	11
2.3. Proceso $t\bar{t} + DM$	11
2.4. Proceso <i>single top</i> +DM	15
2.5. Objetivos del trabajo	20
3. Redes Neuronales Artificiales	21
3.1. Breve contexto histórico	21
3.2. Introducción a las redes neuronales artificiales	22
3.3. TMVA	23
4. Resultados	25
4.1. Desempeño de la red en sucesos $t\bar{t}$ +DM	25
4.2. Predicción en sucesos <i>single top</i> +DM	28
4.3. Comparación con el método de las elipses	32
5. Conclusiones	35

Capítulo 1

Introducción

Una de las mayores ambiciones de la Física en la actualidad es la búsqueda y detección de Materia Oscura. La Materia Oscura (DM por sus siglas en inglés *Dark Matter*) es un tipo diferente de materia que conforma aproximadamente el 27% de la densidad total del Universo. A diferencia de la materia normal, la Materia Oscura no interacciona con la fuerza electromagnética, es decir, no emite, refleja ni absorbe luz, lo cual la hace extremadamente difícil de detectar. Sin embargo, se sabe de su existencia porque presenta efectos gravitacionales en la materia normal.

1.1. Historia y evidencias de Materia Oscura

A lo largo de la historia de la Física existen numerosas evidencias, basadas principalmente en observaciones astronómicas y cosmológicas, cuya fenomenología solo se puede explicar mediante la existencia de un tipo desconocido e invisible de materia, la Materia Oscura. El primer intento de introducir la Materia Oscura para explicar un fenómeno cosmológico vino en 1906 de la mano de Lord Kelvin [1], el cual trató de estimar la cantidad de Materia Oscura presente en la Vía Láctea. Para ello, consideró las estrellas de la galaxia como un gas de partículas bajo la influencia de la gravedad, y aplicando la teoría de los gases dedujo una relación entre el tamaño del sistema y la velocidad de dispersión de las estrellas. Con ello argumentó que la presencia de Materia Oscura ("*matière obscure*", en el artículo original) era similar o ligeramente menor a la cantidad de materia normal.

Más adelante, en 1933, el astrónomo suizo-americano Frank Zwicky estudió varias galaxias en el cúmulo Coma [2], las cuales presentaban una gran diferencia en la velocidad de dispersión con respecto a otros cúmulos de galaxias. Zwicky aplicó el Teorema del virial para obtener la masa del cúmulo y concluyó que la presencia de Materia Oscura es muy superior a la de materia luminosa.

Otra de las grandes evidencias de la existencia de Materia Oscura proviene del estudio de las curvas de rotación de las galaxias, que representan la velocidad orbital del material galáctico en función de su distancia al centro de la galaxia. En la década de 1970, con las mejoras en la resolución y la calidad de las observaciones astronómicas, se pudieron estudiar las curvas de rotación de numerosas galaxias. Se observó que, contrariamente a las

predicciones de la dinámica newtoniana, la velocidad orbital de las estrellas se estabilizaba y se mantenía constante a partir de cierta distancia al centro (Figura 1.1). Esto indicaba que las estrellas exteriores rotaban a una velocidad mayor de la esperada, por lo que debía existir una gran cantidad de masa no visible en el medio galáctico.

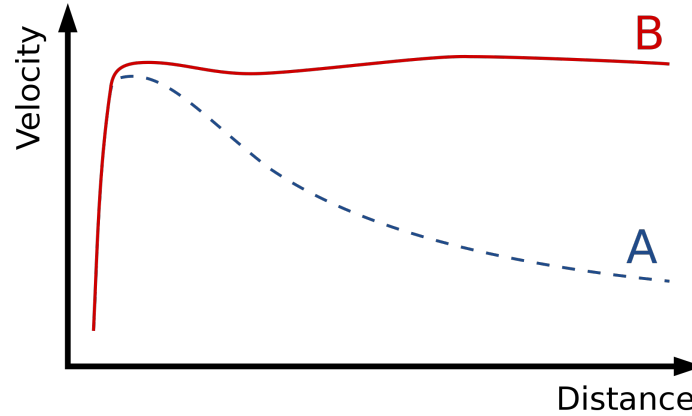


Figura 1.1: Curva de rotación de una galaxia espiral. (A) corresponde a la predicción teórica y (B) a las observaciones.

Otras observaciones que indican la existencia de la Materia Oscura tienen que ver con la confección de mapas de masa usando el mecanismo de lente gravitacional. Este fenómeno consiste en la deformación observada en las imágenes obtenidas de imágenes muy lejanas, debido a que su luz se curva al atravesar campos gravitatorio muy intensos. El estudio detallado de dichas deformaciones permite estimar la localización de la masa de sistemas como clústeres de galaxias en colisión revelando la presencia de altas concentraciones de masa en lugares aparentemente vacíos.

Finalmente, la presencia de la Materia Oscura puede inferirse también a través de su efecto en las anisotropías de temperatura en el fondo cósmico de microondas, permitiendo estimar que aproximadamente el 26 % del total de la materia del Universo es Materia Oscura. [3]

1.2. El Modelo Estándar de la Física de Partículas

La Física de Partículas tiene como objetivo fundamental dar una explicación a las leyes fundamentales de la naturaleza, desde el estudio de las partículas elementales que componen la materia y sus interacciones. La teoría por excelencia a día de hoy que engloba casi todo el conocimiento actual de la Física de Partículas es el Modelo Estándar. Esta teoría fue desarrollada en la década de 1970 [4] y engloba 3 de las 4 interacciones fundamentales de la naturaleza: interacción electromagnética, interacción débil e interacción fuerte (la interacción gravitatoria no se enmarca en el Modelo Estándar).

En el Modelo Estándar, la materia que conocemos está conformada por un tipo de partículas denominadas fermiones, y las interacciones entre los fermiones son mediadas por unas partículas llamadas bosones. Los fermiones son partículas de espín semi-entero

mientras que los bosones tienen espín entero.

Entre los fermiones, podemos distinguir dos grandes grupos: leptones y quarks. A su vez cada uno de estos grupos se divide en tres generaciones de partículas: primera generación, en la que se encuentra el electrón, el neutrino electrónico y los quarks up y down; segunda generación, formada por el muón, neutrino muónico y los quarks charm y strange; y tercera generación, con el tau, el neutrino tauónico y los quarks bottom y top (Figura 1.2). La diferencia esencial entre leptones y quarks es que los leptones no son afectados por la interacción fuerte (QCD), mientras que los quarks presentan carga de color e interaccionan fuertemente.

	Leptons				Quarks			
	Particle	Q	mass/GeV		Particle	Q	mass/GeV	
First generation	electron (e^-)	-1	0.0005		down (d)	-1/3	0.003	
	neutrino (ν_e)	0	$< 10^{-9}$		up (u)	+2/3	0.005	
Second generation	muon (μ^-)	-1	0.106		strange (s)	-1/3	0.1	
	neutrino (ν_μ)	0	$< 10^{-9}$		charm (c)	+2/3	1.3	
Third generation	tau (τ^-)	-1	1.78		bottom (b)	-1/3	4.5	
	neutrino (ν_τ)	0	$< 10^{-9}$		top (t)	+2/3	174	

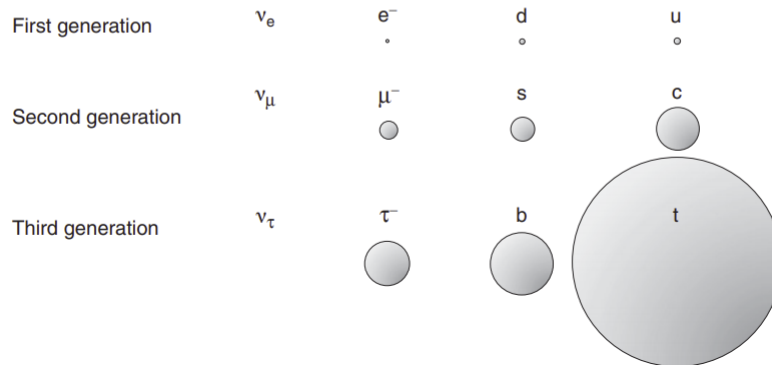


Figura 1.2: Resumen de los fermiones del Modelo Estándar. [5]

La Figura 1.2 muestra las masas relativas de las diferentes partículas. Existe una gran diferencia entre las masas de los fermiones: mientras que los neutrinos son prácticamente no masivos, la masa de los quarks de tercera generación es muy grande, especialmente la del quark top. Este hecho será relevante más adelante en lo que se refiere a búsquedas de Materia Oscura.

En cuanto a las interacciones fundamentales entre fermiones, en el Modelo Estándar cada una de ellas se explica mediante una teoría cuántica de campos (QFT). En las QFT, las interacciones se producen por mediación de unas partículas llamadas bosones (Figura 1.3), que transfieren momento y energía entre los campos fermiónicos. En este sentido, la interacción electromagnética se explica a nivel fundamental mediante la Electrodinámica Cuántica (QED), en la que las partículas con carga intercambian fotones virtuales. Por su parte, la interacción fuerte mantiene los protones y neutrones unidos en el núcleo atómico, y se explica mediante la Cromodinámica Cuántica (QCD) y el intercambio de gluones.

(no masivos). Finalmente, la interacción débil es mediada por los bosones W^+ , W^- y Z (masivos).

El Modelo Estándar se cierra finalmente con la presencia de una nueva partícula, conocida como el bosón de Higgs, responsable de la masa del resto de partículas. El bosón de Higgs fue descubierto en el año 2012 en el Gran Colisionador de Hadrones en el CERN y su masa está en torno a los 125 GeV.

Force	Strength	Boson		Spin	Mass/GeV
Strong	1	Gluon	g	1	0
Electromagnetism	10^{-3}	Photon	γ	1	0
Weak	10^{-8}	W boson	W^\pm	1	80.4
		Z boson	Z	1	91.2

Figura 1.3: Resumen de los bosones del Modelo Estándar. [5]

1.3. La Materia Oscura en Física de Partículas.

El Modelo Estándar de la Física de Partículas, previamente descrito, es una de las teorías físicas más precisas de las que disponemos, y ha demostrado ser muy eficaz a la hora de explicar los experimentos realizados hasta la fecha. Por este motivo, es razonable pensar que la Materia Oscura pueda ser descrita en términos de partículas elementales. Esto es, podemos describir la Materia Oscura mediante una partícula.

No obstante, cabe destacar que aún no existe ningún tipo de evidencia directa de que la Materia Oscura esté compuesta por partículas. De hecho, existen modelos alternativos de Materia Oscura en la que esta se explica mediante objetos macroscópicos (agujeros negros primordiales) o incluso mediante correcciones en modelos gravitatorios (MOND). La hipótesis de una Materia Oscura formada por partículas es la más popular hoy en día. En este trabajo se considera un modelo de Materia Oscura con partículas de tipo no bariónico.

Uno de los requisitos fundamentales para estas partículas es que no posean carga, pues en tal caso interactuarían electromagnéticamente y serían visibles. Además, podemos suponer que se trata de partículas muy masivas, ya que no han aparecido anteriormente en experimentos de bajas y medias energías. Por último, estas partículas han de tener un acoplamiento pequeño con las partículas del Modelo Estándar, ya que se sabe que interactúan poco con la materia ordinaria.

Con todo lo expuesto, para este trabajo se ha considerado un modelo de Materia Oscura con una partícula mediadora ϕ masiva y eléctricamente neutra. Según este modelo, esta partícula mediadora puede ser de naturaleza escalar (espín 0) o pseudoescalar (espín 0, pero cambia de signo bajo transformaciones de paridad), se acoplaría a un par $t\bar{t}$ y decaería a dos fermiones de Materia Oscura χ . En el contexto de este modelo, se han estudiado dos posibles procesos que se pueden dar en un colisionador de hadrones, como el experimento CMS. Estos serán explicados en detalle en el Capítulo 2.

1.4. Experimento CMS

El LHC (*Large Hadron Collider*) [6] es el mayor acelerador de partículas del mundo. Fue construido en el CERN (Organización Europea para la Investigación Nuclear) entre 1998 y 2008, y posteriormente actualizado para lograr energías de colisión de hasta 13 TeV. Se encuentra situado en un túnel circular de 27 km de circunferencia, entre la frontera de Francia y Suiza y en él se realizan los experimentos de física de partículas a mayor energía del mundo. En el LHC se aceleran haces de protones (*bunches*) hasta energías relativistas (~ 13 TeV) y se hacen colisionar en 4 puntos estratégicos, dotados de detectores que recogen información acerca de las colisiones. Estos detectores son: LHCb, ATLAS, ALICE y CMS.

Particularmente, el experimento CMS (*Compact Muon Solenoid*) [7] (Figura 1.4) consta de un enorme solenoide de 21 metros de largo por 16 de ancho especializado en detección de muones. Está formado por 4 capas de detectores y un imán solenoidal dispuestos de modo cilíndrico alrededor de la zona de colisión de los protones. Cada una de estas capas recoge diferente información acerca de los productos de las colisiones [8].

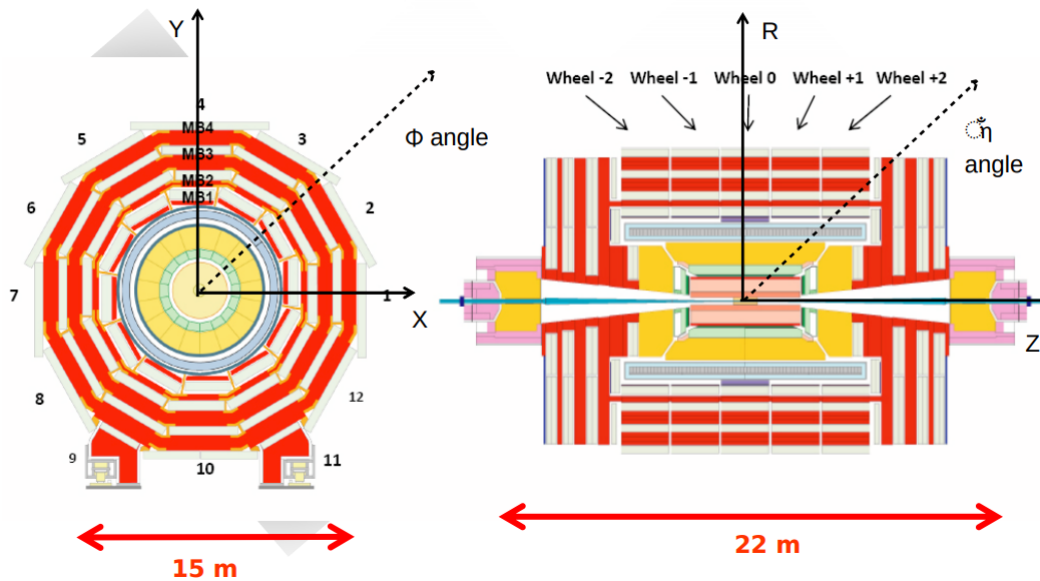


Figura 1.4: Esquema de la estructura del experimento CMS.

La capa más interna de todas es el detector de trazas (*tracker*), que rodea el punto de interacción. Consta de una alta densidad de píxeles de silicio que permiten medir el momento y la trayectoria de partículas cargadas, así como detectar dónde decaen algunas partículas inestables de vida media alta. Alrededor de ella se sitúa el calorímetro electromagnético, que permite medir con precisión la energía de fotones y electrones, deteniéndolos completamente. Seguidamente está el calorímetro hadrónico, que de igual forma mide la energía de las partículas hadrónicas que lo atraviesan. Exterior a los calorímetros se encuentra el imán, que ayuda a curvar las partículas y medir su relación carga/masa.

Por último, la capa más exterior de todas es el detector de muones, que mide la trayectoria y el momento de los muones.

El detector CMS es capaz de procesar las señales electrónicas dejadas por las partículas producidas en las colisiones en los diferentes subdetectores, para identificar y reconstruir con gran precisión su cuadrimomento. A continuación se explican brevemente las variables más relevantes con las que se ha trabajado en este proyecto.

En primer lugar están las variables cinemáticas habituales: masa invariante (m) y momento transverso (p_T), el cual hace referencia a la componente del momento que es perpendicular a la trayectoria de los haces de partículas.

Por otra parte, se recogen dos variables relacionadas con la trayectoria de la partícula. La primera de ellas es el ángulo ϕ , que se define como el ángulo de la trayectoria en el plano transverso. La otra variable es la pseudorrapidez η . Se define como $\eta = -\log(\tan(\theta/2))$ donde θ es el ángulo entre el eje Z y la trayectoria de la partícula. η mide el ángulo entre la trayectoria de la partícula y el eje Z, siendo $\eta = 0$ la dirección perpendicular al eje Z y $\eta = \infty$ el eje Z. Por razones de construcción CMS solo puede detectar partículas con $0 \leq \eta \leq 2,4$.

Cabe mencionar que algunas partículas como los neutrinos no son recogidas por los detectores de CMS, por lo que atraviesan el detector sin dejar traza. Sin embargo, se puede tratar de estimar su energía, para lo cual se define la variable MET (*Missing Energy Transverse*). Como los haces de protones inicialmente solo llevan momento en el eje Z, por el principio de conservación, el momento total final en el plano transverso de todas las partículas producto de la colisión ha de ser igual a 0. De esta forma, la MET se define como

$$\vec{MET} = - \sum_{\substack{\text{partículas} \\ \text{visibles}}} \vec{p}_T$$

1.5. Búsquedas de Materia Oscura

En la actualidad existen multitud de experimentos cuyo objetivo es la detección y el estudio de la Materia Oscura. Estos se pueden clasificar en tres grupos: búsqueda directa, búsqueda indirecta y búsqueda en colisionadores. Los experimentos de búsqueda directa están orientados a detectar trazas de partículas o núcleos que han interactuado con Materia Oscura a su paso por la Tierra. Estos experimentos constan de un gran detector situado a mucha profundidad, para minimizar el ruido de otras fuentes como los rayos cósmicos. Las búsquedas indirectas por su parte se enfocan a encontrar restos de Materia Oscura que interactúa o decae en zonas del espacio exterior donde la densidad es alta. En concreto, se buscan fotones muy energéticos que emanan de la galaxia. Por último, las búsquedas en colisionadores de partículas tratan de producir las partículas que conforman la Materia Oscura mediante colisiones de protones generalmente, y estudiar los productos de estas colisiones. En este caso, se espera que la Materia Oscura sea detectada como una gran cantidad de energía o momento faltante.

En concreto este trabajo se relaciona con este último grupo de experimentos. Se enmarca en una búsqueda de Materia Oscura a partir de las colisiones protón-protón que

se llevan a cabo en el experimento CMS. En concreto se van a investigar dos procesos de producción de Materia Oscura: en asociación con un par quark top - quark antitop, y a partir de un solo quark top que decae leptónicamente.

Capítulo 2

Procesos de producción de Materia Oscura con quarks top

Este trabajo se centra en el estudio de procesos de producción de Materia Oscura en asociación con quarks top. Previamente se ha explicado que las partículas candidatas a conformar la Materia Oscura se espera que sean muy masivas. Esto conlleva una interacción muy fuerte con el bosón de Higgs, que es la partícula del SM responsable de la masa del resto de partículas. Por esta razón, es razonable pensar que existe una relación entre las partículas de Materia Oscura y la partícula más masiva del Modelo Estándar, el quark top.

En esete sentido, en este Capítulo se introduce la interacción top - antitop en el Modelo Estándar, así como dos procesos de producción de DM en asociación con quarks top.

2.1. Simulación de sucesos

El estudio de los procesos de producción de DM presentado en este trabajo se basa en el análisis de una gran cantidad de sucesos. Estos han sido generados mediante simulaciones de Monte Carlo, con el simulador MadGraph5_aMC@NLO [9]. Este programa aplica las reglas de Feynmann para un proceso determinado obteniendo su sección eficaz diferencial para a continuación generar sucesos recorriendo aleatoriamente dichas secciones eficaces. MadGraph5_aMC@NLO es capaz de calcular dichas secciones eficaces con precisión NLO (*Next to Leading Order*), tanto para procesos del Modelo Estándar, como para procesos más allá del Modelo Estándar, siempre que sean formulados en el contexto de una teoría cuántica de campos. Para el estudio realizado en este trabajo, no se ha llevado a cabo la simulación detallada de las partículas generadas a lo largo del detector.

2.2. Interacción top - antitop estándar ($t\bar{t}$)

En esta primera sección se introduce el proceso top - antitop del Modelo Estándar, que es el punto de partida de los dos procesos de Materia Oscura que se estudian en este trabajo.

Como se ha mencionado en la Sección 1.2, el quark top es el fermión más masivo del

Modelo Estándar, siendo su masa más dos órdenes de magnitud superior a la masa del protón. Esta característica hace que el top decaiga muy rápidamente ($5 \cdot 10^{-25}$ s) y no llegue a hadronizar (generación de un jet de partículas debido a las propiedades de la fuerza fuerte).

La única forma conocida de decaimiento del quark top es a través de la interacción débil, produciendo un bosón W^+ más un quark (b, s, d). En este trabajo se estudiará el decaimiento más probable, que es $t \rightarrow W^+ b$, con un branching ratio de

$$\Gamma(W^+ b)/\Gamma(W^+ q) = (91 \pm 4) \%$$

Tras esta desintegración, el bosón W^+ decaerá. El canal de desintegración que se va analizar es $W^+ \rightarrow l^+ \nu_l$. Esto es porque al decaer en un leptón cargado y un neutrino, la señal que deja el suceso en el detector es limpia. Puesto que la probabilidad de decaimiento en $e^+ \nu_e$, $\mu^+ \nu_\mu$, $\tau^+ \nu_\tau$ es aproximadamente la misma, no es necesario distinguir entre estos tres modos. El branching ratio de este decaimiento¹ es

$$\Gamma(l^+ \nu_l)/\Gamma_{total} = (10,86 \pm 0,09) \%$$

A su vez, el quark antitop sufre un decaimiento análogo con las cargas opuestas: $\bar{t} \rightarrow l^- \bar{\nu}_l \bar{b}$. Se denotarán como decaimientos $t\bar{t}$ estándar a los de la forma $t\bar{t} \rightarrow b\bar{b} l^+ \nu_l l^- \bar{\nu}_l$, para diferenciarlos de aquellos que involucren Materia Oscura.

El canal de producción más importante de pares de quarks top en el LHC es el conocido como fusión de gluones, en la que dos gluones interactúan entre sí, dando lugar a un gluón muy energético que puede originar el par top - anti top el cual decaerá como se ha descrito anteriormente. En la Figura 2.1 se puede observar el diagrama de Feynman del proceso $t\bar{t}$ estándar.

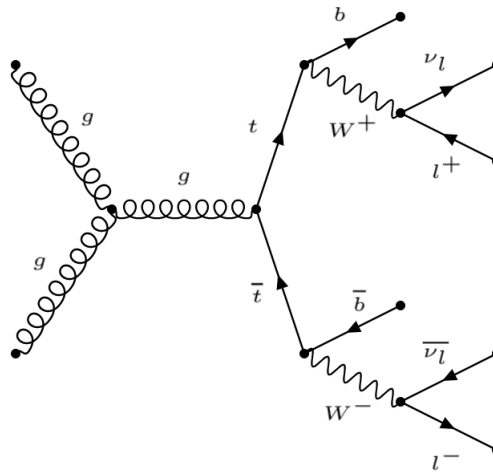


Figura 2.1: Diagrama de Feynman del proceso $t\bar{t}$ estándar.

¹Branching ratio medio para los tres decaimientos (e, μ, τ).

2.2.1. Reconstrucción del momento de los neutrinos en el proceso $t\bar{t}$ estándar

Los neutrinos son partículas que no dejan ningún rastro o traza en los detectores, por lo que no puede medirse directamente su momento. Sin embargo, para el proceso $t\bar{t}$ estándar se puede reconstruir dicho momento a partir de las restricciones de la masa invariante de los quarks top, los bosones W y de la MET ([10], [11]).

2.3. Proceso $t\bar{t} + DM$

En el proceso $t\bar{t}+DM$ se producen dos pares top - antitop a partir de dos gluones muy energéticos. Posteriormente un quark top y un antitop se aniquilan y dan lugar a un mediador ϕ que decae en dos partículas de Materia Oscura χ . Por su parte, el otro quark top y el antitop decaen de igual forma que en el proceso $t\bar{t}$ estándar. En la Figura 2.2 se presenta el diagrama de Feynman del proceso.

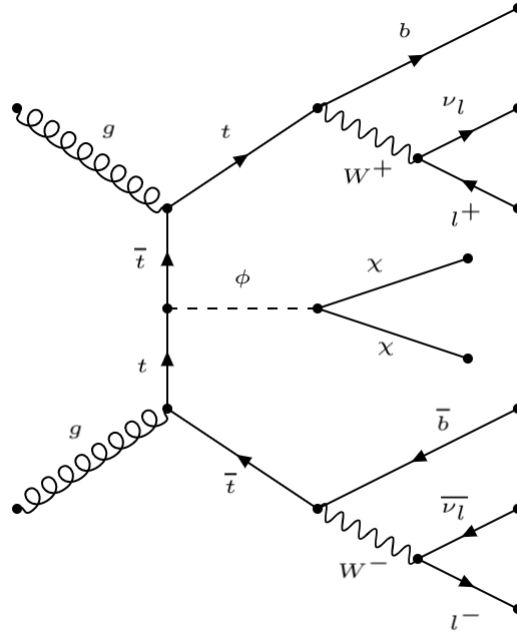


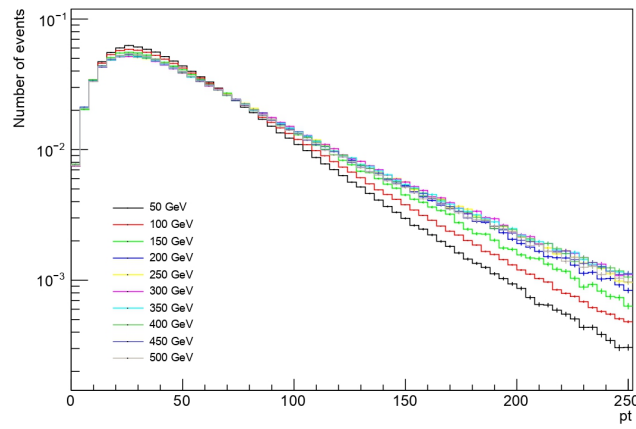
Figura 2.2: Diagrama de Feynman del proceso $t\bar{t} + DM$.

Si bien para el proceso $t\bar{t}$ estándar se puede reconstruir de forma analítica el momento de los neutrinos, este no es el caso para el proceso $t\bar{t}+DM$. La razón última de ello es que la introducción del mediador ϕ añade nuevas variables al sistema pero el número de restricciones sigue siendo el mismo que en el proceso $t\bar{t}$ estándar.

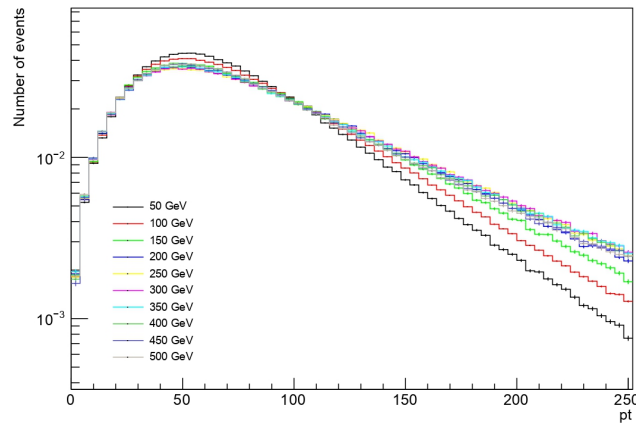
Sin embargo, existen formas de reconstruir el momento del mediador de Materia Oscura a partir de las restricciones del proceso $t\bar{t}$ estándar. Este fue el objeto de estudio de Lantero A. [12]. De igual forma, el primer objetivo del trabajo aquí presentado es resolver este problema haciendo uso de las redes neuronales artificiales.

Cuando se estudia este proceso en un detector como CMS, se recogerán únicamente los productos finales de la colisión. En primer lugar se obtiene información sobre el momento y la trayectoria de los dos leptones. Por otra parte, los quarks bottom hadronizarán² y se recogerá información sobre ellos reconstruida a partir de los jets de hadrones. Finalmente, se recogerá la MET, que contiene la información sobre el momento de las partículas que escapan a los detectores: neutrinos y Materia Oscura.

En las posteriores Figuras se muestran las distribuciones de algunas variables cinemáticas que caracterizan al proceso $t\bar{t}$ +DM. Los histogramas están separados para cada masa del mediador ϕ .

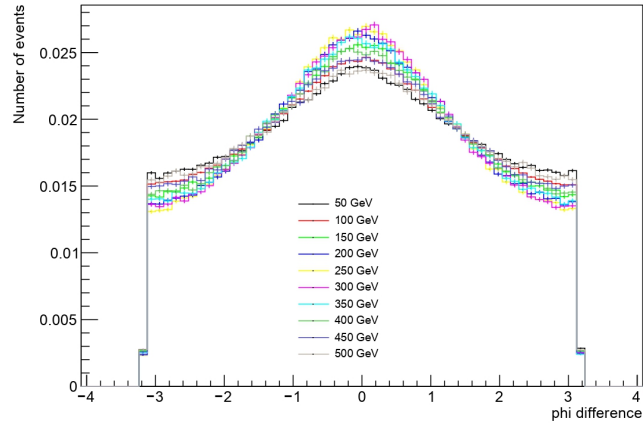
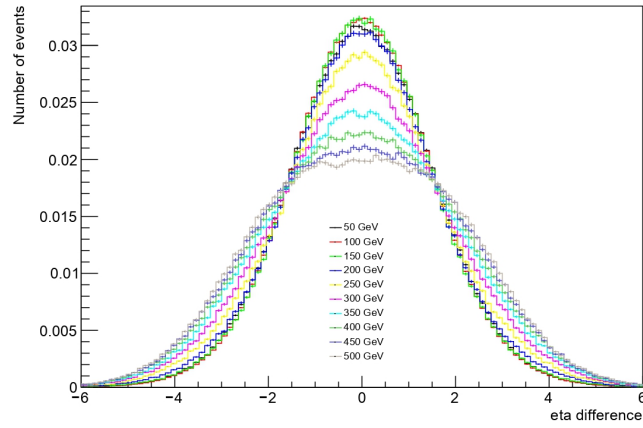
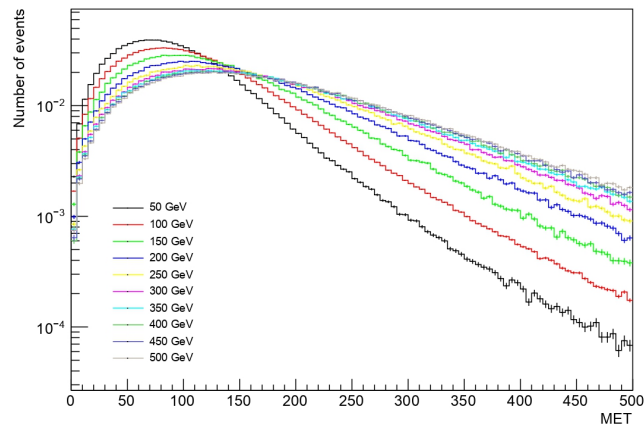


(a) Distribución de momento de los leptones.



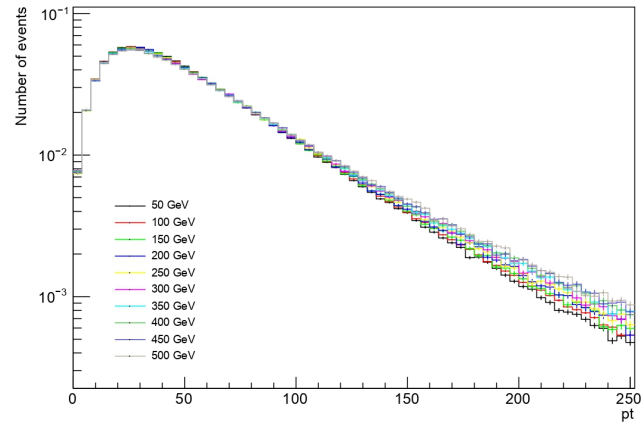
(b) Distribución de momento de los quark bottom.

²Cuando se libera un solo quark, se generan en el vacío pares quark-antiquark que se combinan y forman chorros de hadrones (jets). Este proceso se llama hadronización.

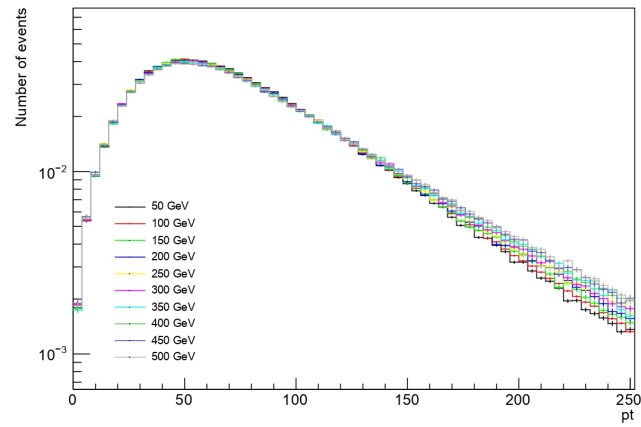
(c) Correlaciones del ángulo ϕ .(d) Correlaciones de la variable η .

(e) Distribución de la MET.

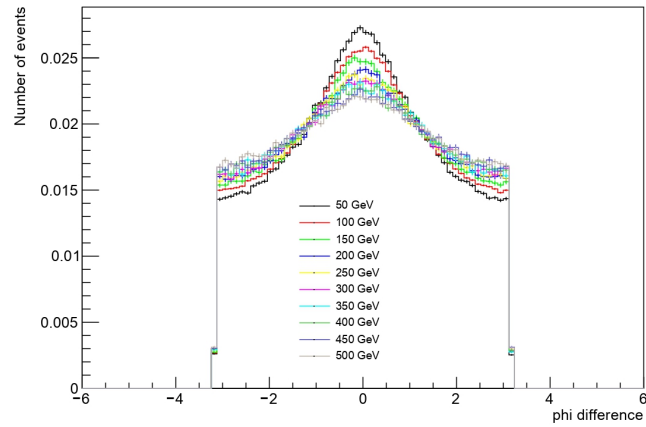
Figura 2.3: Variables cinemáticas que caracterizan al proceso $t\bar{t} + DM$ con mediador escalar. Las subfiguras (a), (b) y (e) están en escala logarítmica.

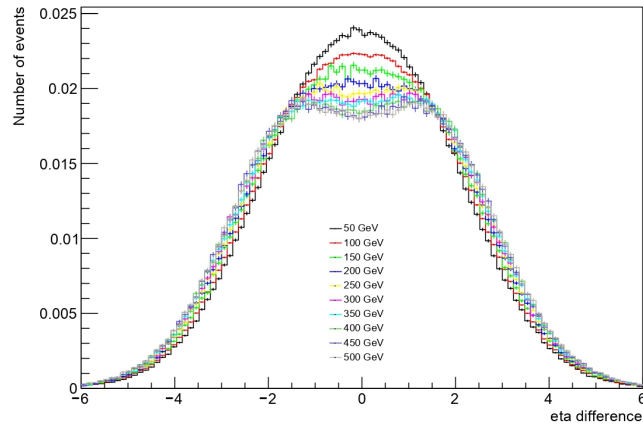
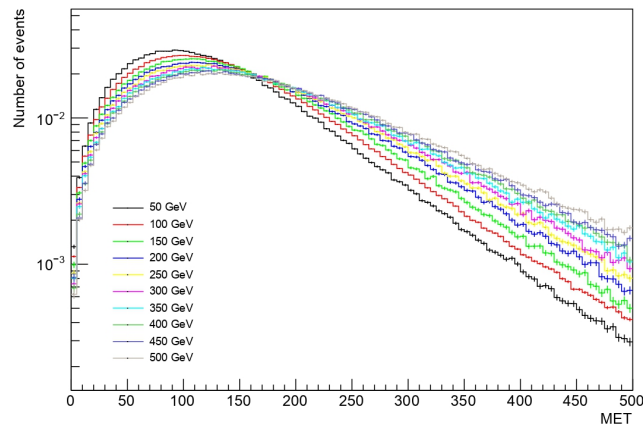


(a) Distribución de momento de los leptones.



(b) Distribución de momento de los quark bottom.

(c) Correlaciones del ángulo ϕ .

(d) Correlaciones de la variable η .

(e) Distribución de la MET.

Figura 2.4: Variables cinemáticas que caracterizan al proceso $t\bar{t}$ +DM con mediador pseudoescalar. Las subfiguras (a), (b) y (e) están en escala logarítmica.

2.4. Proceso *single top*+DM

En el proceso *single top*+DM también se genera un par top - antitop que da lugar a una partícula ϕ de Materia Oscura. Sin embargo, solo uno de los tops proviene de un gluon energético que decae a un par $t\bar{t}$. El otro top proviene de un quark bottom que decae débilmente a un quark top y un bosón W . En la Figura 2.5 se muestra el diagrama de Feynman del proceso.

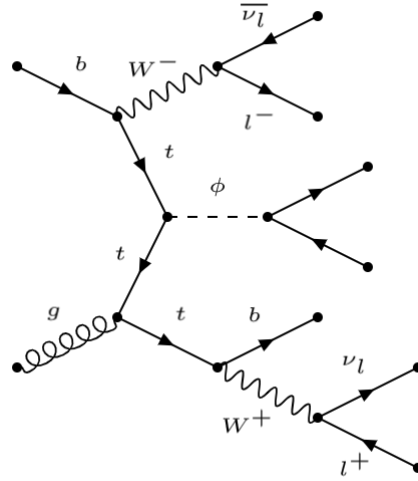
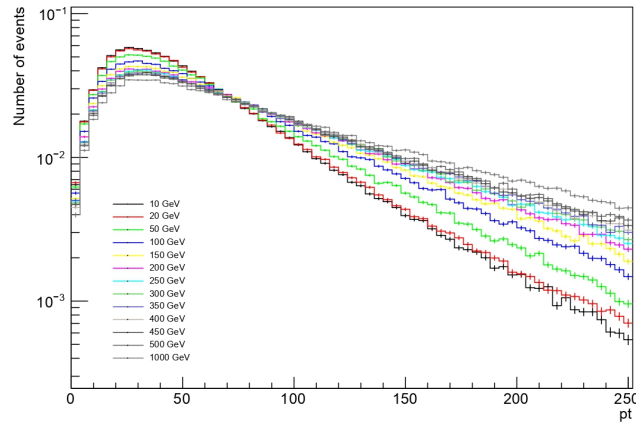


Figura 2.5: Diagrama de Feynman del proceso *single top*+DM.

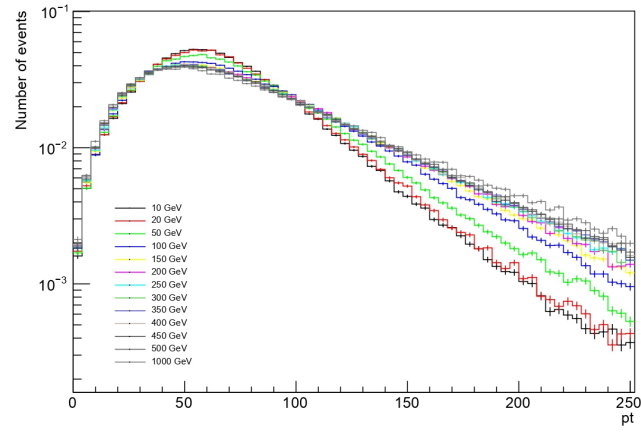
A diferencia del proceso $t\bar{t}$ +DM, en el proceso *single top*+DM no se puede constreñir el momento de los neutrinos, por lo que no existe ningún método analítico para reconstruir el momento del mediador ϕ .

Otra diferencia importante con el proceso anterior es la información que disponemos cuando estudiamos un proceso *single top*+DM con un detector. Al igual que antes, se recoge la misma información sobre los dos leptones y la MET. Pero en este caso, solo se recoge un quark bottom (en forma de jet), por lo que la información es más escasa.

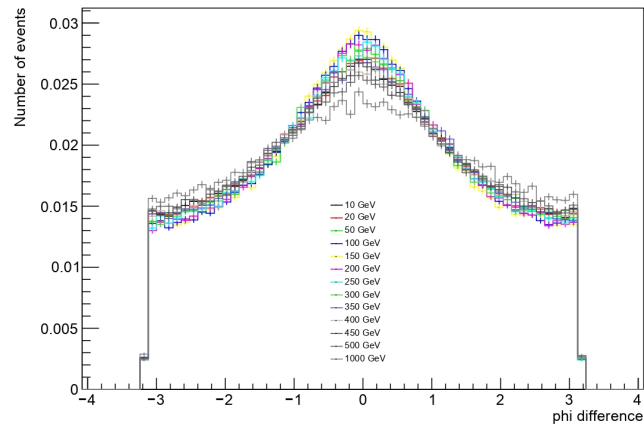
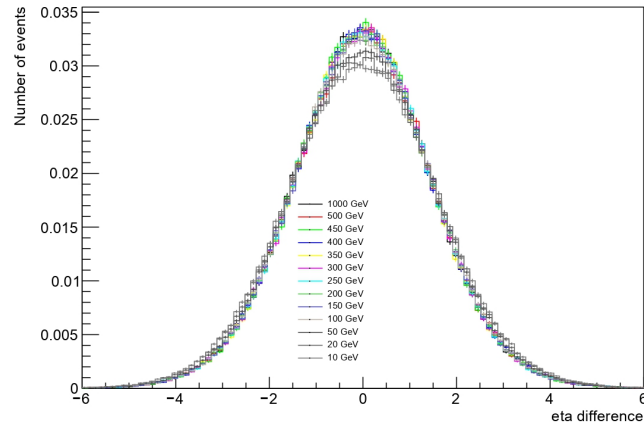
En las posteriores Figuras se presentan algunas variables cinemáticas relevantes en el proceso *single top*+DM.

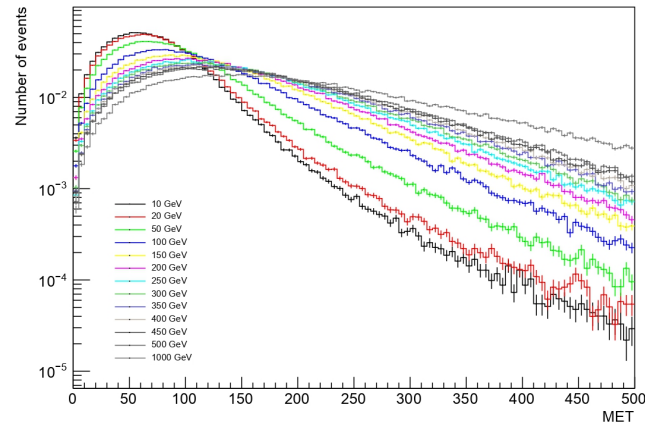


(a) Distribución de momento de los leptones.



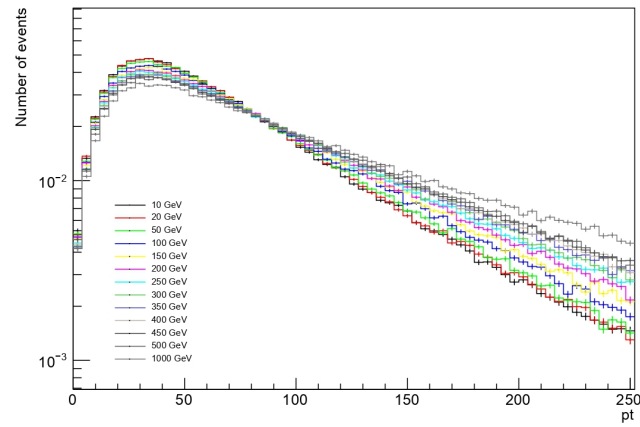
(b) Distribución de momento del quark bottom.

(c) Correlaciones del ángulo ϕ .(d) Correlaciones de la variable η .

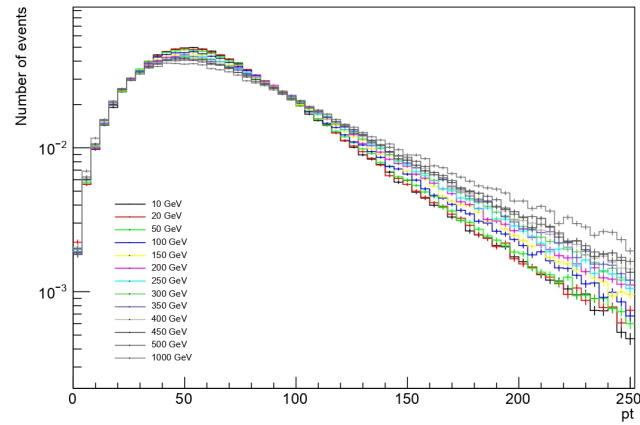


(e) Distribución de la MET

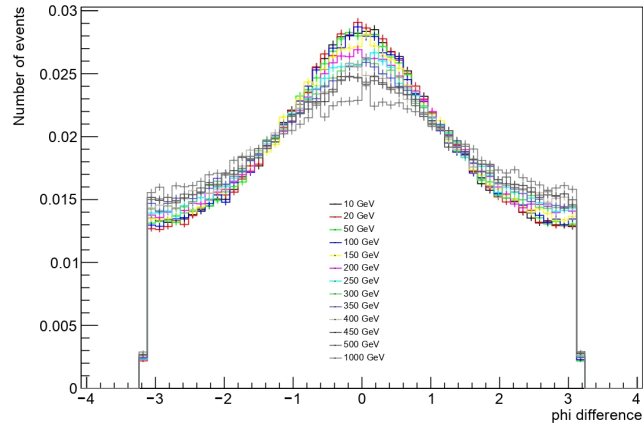
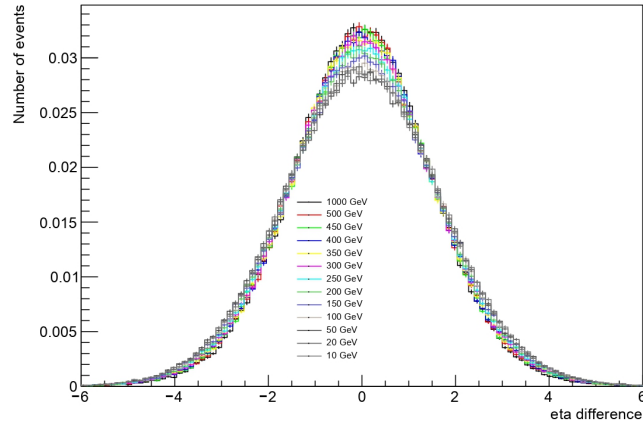
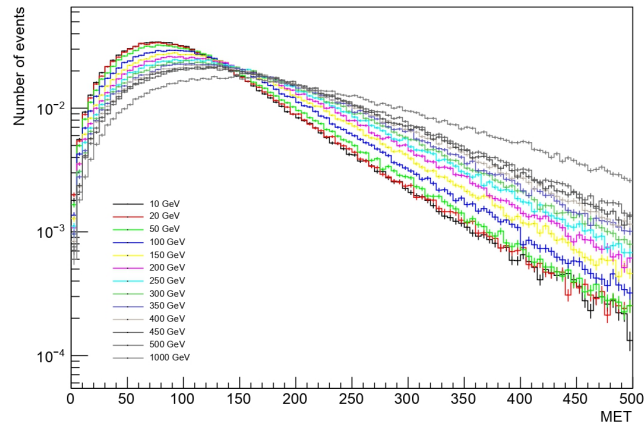
Figura 2.6: Variables cinemáticas que caracterizan al proceso single top + DM con mediador escalar. Las subfiguras (a), (b) y (e) están en escala logarítmica.



(a) Distribución de momento de los leptones.



(b) Distribución de momento del quark bottom.

(c) Correlaciones del ángulo ϕ .(d) Correlaciones del ángulo η .

(e) Distribución de la MET.

Figura 2.7: Variables cinemáticas que caracterizan al proceso single top + DM con mediador pseudoescalar. Las subfiguras (a), (b) y (e) están en escala logarítmica.

2.5. Objetivos del trabajo

Una vez presentados los procesos de producción de Materia Oscura que se van a estudiar, de ahora en adelante es importante fijar los objetivos de este trabajo. Esto es en primer lugar para tener una visión global de lo que se pretende conseguir, y en segundo lugar para entender el porqué y la relevancia de los resultados que se presentan en el Capítulo 4.

El primero de los objetivos es comprobar si una red neuronal regresiva es capaz de predecir el momento del mediador de Materia Oscura ϕ en sucesos $t\bar{t}$ +DM y, en tal caso, comparar la calidad del método con los métodos analíticos de la literatura.

El segundo objetivo y el más interesante de este estudio consiste en analizar si una red neuronal en modo regresión es capaz de estimar el momento del mediador ϕ en sucesos *single top*+DM. Esto es importante porque a diferencia del proceso $t\bar{t}$ +DM, para el *single top*+DM no existen métodos analíticos. Por tanto, de resultar satisfactorio este sería el primer método desarrollado para reconstruir el momento del ϕ en sucesos *single top*+DM.

Es importante destacar también que el momento del mediador ϕ tiene gran interés para discriminar sucesos de Materia Oscura frente a sucesos del SM, y también para caracterizar cinemáticamente sucesos de Materia Oscura.

Capítulo 3

Redes Neuronales Artificiales

Las redes neuronales artificiales (*artificial neural networks*, ANN) son modelos computacionales diseñados para modelar problemas no-lineales. Reciben este nombre porque toman como fuente de inspiración las redes neuronales biológicas presentes en el cerebro de los animales. Consisten en una serie de nodos interconectados llamados 'neuronas', que imitan el comportamiento de las neuronas biológicas. Cada conexión, al igual que la sinapsis en el cerebro, transmite información a las neuronas de alrededor. Las neuronas que forman la red reciben unas señales de entrada (*inputs*) en forma de números reales y computan la señal de salida (*output*) a partir de una función no lineal de la suma de sus *inputs*. Típicamente estas neuronas se agrupan en capas, y cada capa realiza transformaciones diferentes de los *inputs* que recibe. Cada neurona y cada conexión tienen asociado un peso dentro de la red, el cual se va modificando a medida que avanza el aprendizaje.

Este tipo de modelo es el más popular hoy en día en el contexto del aprendizaje automático, empleado a día de hoy en numerosas tareas como visión artificial, reconocimiento de voz o procesamiento de datos. En este trabajo en particular las redes neuronales artificiales se aplican con el objetivo de realizar análisis de regresión y modelado de datos.

3.1. Breve contexto histórico

El primer intento de realizar un modelo computacional de redes neuronales fue en 1940 por Warren McCulloch y Walter Pitts. A finales de los años 40, D.O. Hebb ideó una hipótesis de aprendizaje llamada 'aprendizaje Hebbian' basada en el mecanismo de plasticidad neuronal. Más adelante, en 1958 Rosenblatt creó el perceptrón, un algoritmo de aprendizaje supervisado capaz de realizar una clasificación binaria. Sin embargo, las primeras redes multicapa funcionales fueron publicadas en 1965 por Ivakhnenko y Lapa.

Ya en la década de 1970 surgió, de la mano de Werbos, el algoritmo de propagación hacia atrás (*backpropagation*), clave en el desarrollo histórico de las redes neuronales artificiales. A partir de este avance, el desarrollo de las ANN se estancó durante varios años. No fue hasta principios del siglo XXI cuando las redes neuronales artificiales volvieron a cobrar gran importancia. Esto fue debido al desarrollo del CMOS (*complementary metal-oxide-semiconductor*), la mejora de las GPUs, la computación distribuida y el aumento de la capacidad computacional en general, avances que permitieron aplicar de forma muy

eficiente y a gran escala los algoritmos de aprendizaje y las redes neuronales artificiales.

3.2. Introducción a las redes neuronales artificiales

En esta sección se introducen los conceptos básicos que caracterizan la ANN. En primer lugar, toda red neuronal tiene tres componentes fundamentales:

- **Neuronas:** son las unidades básicas de la red. Reciben numerosos *inputs* de las neuronas de alrededor, y producen una señal de salida que transmiten a otras neuronas.
- **Conexiones:** las neuronas están interconectadas. Cada una de estas conexiones recibe un peso, que determina la importancia de dicha conexión dentro de la red. Una misma neurona puede tener múltiples conexiones de entrada y de salida.
- **Función de activación:** consiste en una función que calculan las neuronas a partir de las señales de entrada que reciben, y que les permite obtener el *output*. Típicamente se trata de funciones no lineales, como la sigmoide o la tangente hiperbólica. Habitualmente a estas funciones se les suele incluir un término de bias.

En cuanto a la organización de las ANN, estas suelen disponer de una arquitectura de capas (Figura 3.1). La primera capa de todas corresponde típicamente a las variables de entrada de la red, mientras que la última capa computa los valores de salida. El resto de capas se denominan capas intermedias (*hidden layers*), y realizan transformaciones sucesivas de las variables de entrada hasta obtener los valores de salida.

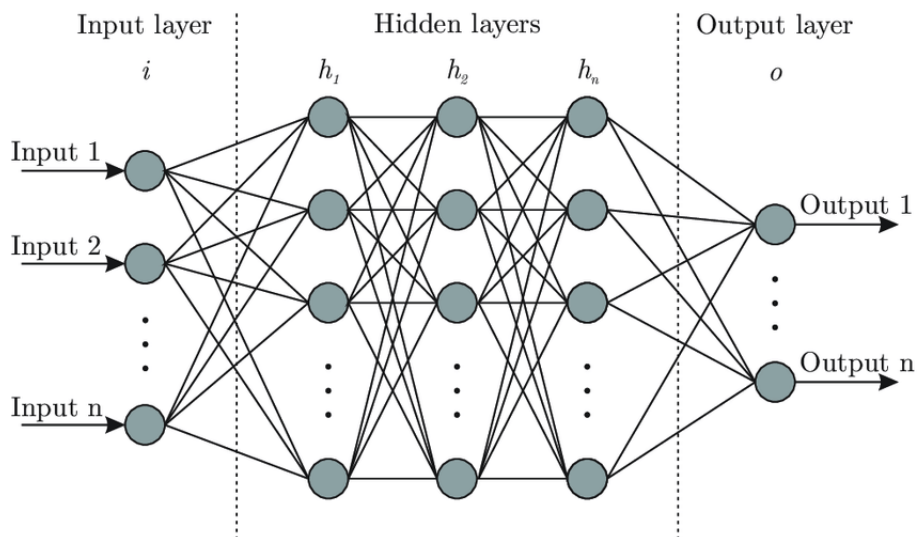


Figura 3.1: Esquema de la arquitectura de una red neuronal artificial. [13]

Una característica muy importante de las ANN, y que es precisamente lo que las hace tan potentes, es su capacidad de aprendizaje. Este aprendizaje consiste en ir ajustando los pesos de la red para reducir el error en el resultado. Existen diversas formas concretas de realizar este procedimiento, pero concretamente en este trabajo se utiliza el aprendizaje

supervisado. En el aprendizaje supervisado, se le proporcionan a la red tanto las variables de entrada como los outputs deseados. La tarea de aprender consiste entonces en ir adaptando los pesos de la red hasta que esta produzca un *output* lo más cercano posible al deseado.

El aprendizaje supervisado consta de los siguientes pasos. En primer lugar, se define una función de coste, que es un estadístico función de los *outputs* de la red y de los *outputs* deseados. Típicamente esta función no es más que el error cuadrático medio. Seguidamente, se aplica un algoritmo para tratar de minimizar la función de coste. El más habitual es el de propagación hacia atrás (*backpropagation*), que calcula el gradiente de la función de coste y actualiza los pesos utilizando por ejemplo descenso del gradiente. Este algoritmo tiene la ventaja de que es muy eficiente porque calcula el gradiente capa por capa hacia atrás desde la última hasta la primera.

En el caso concreto de este trabajo se emplean redes neuronales de aprendizaje supervisado en modo de regresión, esto es, se proporciona a la red las variables cinemáticas involucradas en los procesos explicados anteriormente ($t\bar{t}$ +DM y *single top*+DM) así como el momento real que lleva la partícula mediadora de materia oscura ϕ . Con ello, para cada suceso la red produce un *output* con su 'predicción' para el momento del ϕ , y aplica el algoritmo de propagación hacia atrás para actualizar los pesos y reducir el error. La arquitectura específica de las redes utilizadas se detalla en el Capítulo 4, junto a la presentación de los resultados.

3.3. TMVA

Hoy en día existen numerosos paquetes y herramientas informáticas que implementan redes neuronales artificiales en múltiples lenguajes de programación (Java, R, MatLab, Python...). En este trabajo se ha empleado el paquete TMVA de análisis multivariante desarrollado para ROOT¹.

TMVA (*Toolkit for Multivariate Data Analysis*) es un paquete desarrollado para ROOT que implementa diversas técnicas de análisis multivariante para resolver problemas de clasificación y regresión, especialmente orientado a aplicaciones en física de altas energías. TMVA, al igual que ROOT, está desarrollado en C++. Sin embargo, posee una implementación en Python (pyROOT) que es con la que se ha trabajado en este proyecto.

¹ROOT es un *framework* de análisis de datos especializado para física de altas energías.

Capítulo 4

Resultados

En este Capítulo se presentan todos los resultados obtenidos en el trabajo, recogidos en dos secciones diferenciadas. La primera de ellas hace referencia al análisis del uso de las ANN para predecir el momento de la partícula ϕ en sucesos $t\bar{t}$ +DM. La segunda de ellas es la aplicación de las ANN a sucesos *single top*+DM, para los cuales no existe ningún método analítico para aproximar el momento del ϕ .

4.1. Desempeño de la red en sucesos $t\bar{t}$ +DM

En esta sección se recogen los resultados orientados a resolver el primero de los objetivos de este trabajo: comprobar si es posible aplicar una red neuronal para la reconstrucción del momento del mediador de Materia Oscura ϕ . Se han distinguido los casos de mediador escalar y pseudoescalar y se ha entrenado una red neuronal mediante aprendizaje supervisado para cada caso.

En ambos casos la arquitectura de la red ha sido la misma: dos capas intermedias de N y $N/2$ neuronas, donde $N = 14$ es el número de variables proporcionadas a la red. Estas variables han sido el momento transverso, las coordenadas η y ϕ de los dos quarks b y los dos leptones que intervienen en el proceso, más la variable MET (valor absoluto y coordenada ϕ). Además, se ha aplicado un corte en el momento de los leptones para seleccionar solo aquellos con momento mayor a 15 GeV. La razón de este corte tiene que ver con el hecho de que el experimento CMS tiene limitaciones para registrar sucesos con momentos transversos bajos para los objetos físicos.

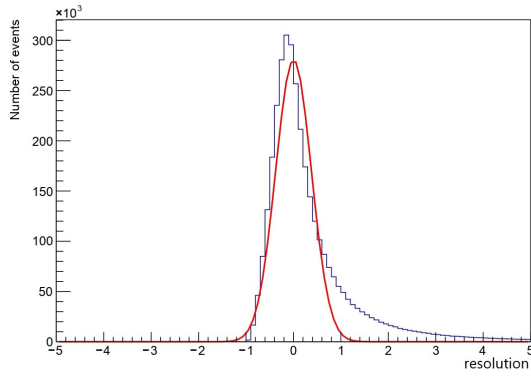
La función de activación de las neuronas ha sido la tangente hiperbólica y el algoritmo de aprendizaje ha sido el de propagación hacia atrás. En los sucesos empleados para entrenar y probar la red, la masa del mediador ϕ varía discretamente entre 50 y 500 GeV, en intervalos de 50 GeV. El número de sucesos disponibles ha sido 8,4 millones en el caso escalar y 6,3 millones para pseudoescalar, de los cuales se han utilizado la mitad para entrenar la red y la otra mitad para probarla.

En la Figura 4.1 se pueden observar los resultados de la resolución de la red para

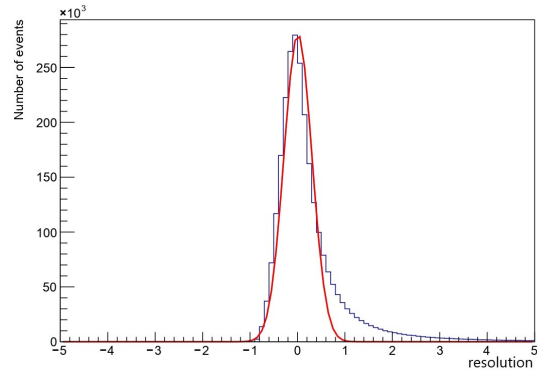
ambos tipos de mediadores. La resolución se define como

$$\text{Resolución} = \frac{p_{\phi,reg} - p_{\phi,true}}{p_{\phi,true}}$$

donde $p_{\phi,reg}$ es el momento reconstruido por la red y $p_{\phi,true}$ es el momento real del mediador.



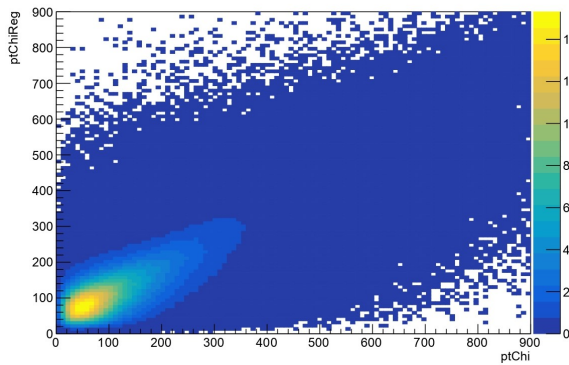
(a) Mediador escalar.
 $\mu_s = (1,3 \pm 0,4) \cdot 10^{-3} \text{ GeV}$
 $\sigma_s = 0,3771 \pm 0,0002 \text{ GeV}$



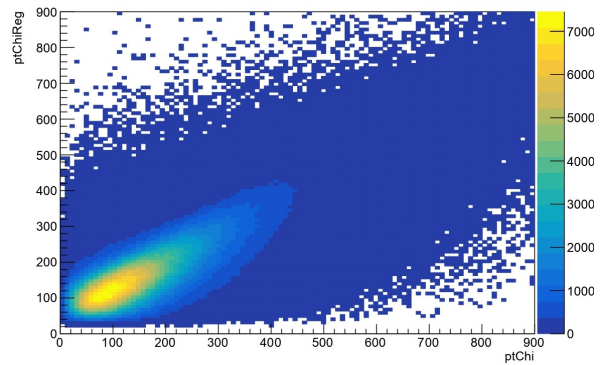
(b) Mediador pseudoescalar.
 $\mu_{ps} = (1,00 \pm 0,03) \cdot 10^{-2} \text{ GeV}$
 $\sigma_{ps} = 0,2945 \pm 0,0001 \text{ GeV}$

Figura 4.1: Histogramas de resolución obtenida por la red neuronal artificial. Los ajustes (rojo) corresponden a dos distribuciones gaussianas, cuya media y desviación se indican arriba.

Por su parte, en la Figura 4.2 se muestran las distribuciones del momento del mediador. En los dos primeros histogramas se recoge para cada suceso el momento real del mediador (eje X) frente al momento reconstruido por la red (eje Y). Los dos últimos muestran superpuestas las distribuciones real y reconstruida del momento del mediador.



(a)



(b)

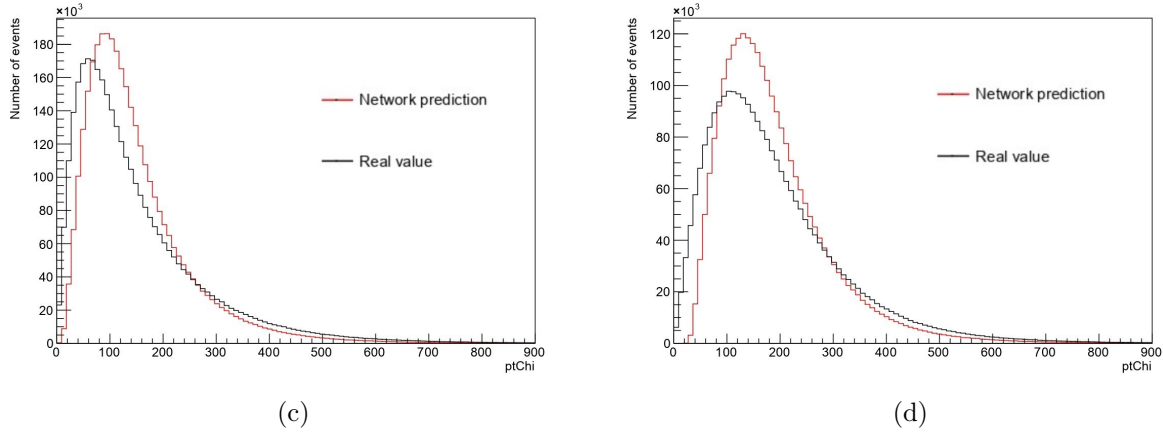


Figura 4.2: Distribución del momento del mediador de Materia Oscura. Los histogramas de la izquierda corresponden a mediador escalar y los de la derecha a mediador pseudoescalar.

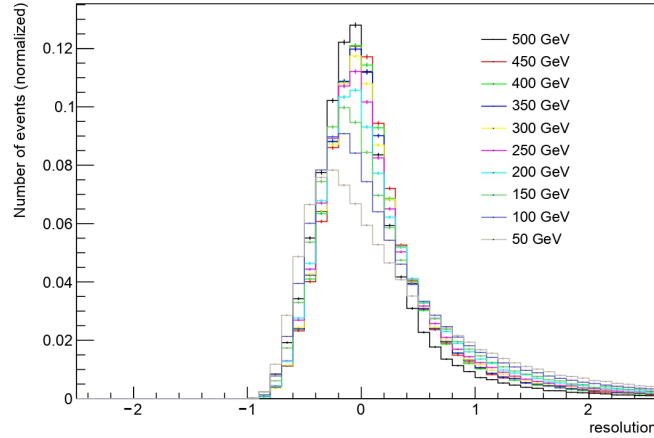
A la vista de la Figuras anteriores, podemos afirmar que las redes neuronales han sido capaces de reconstruir de una forma eficaz el momento del mediador ϕ para sucesos $t\bar{t} + \widetilde{DM}$, tanto con mediador escalar como pseudoescalar.

Los histogramas de resolución muestran que esta se ajusta a una adistribución gaussiana de media muy cercana a 0 y desviación típica $\sim 0,38$ (escalar) y $\sim 0,29$ (pseudoescalar). Esto quiere decir que la red neuronal es capaz de predecir el momento del ϕ en un suceso con un error medio del 38 % y 29 % respectivamente, lo cual es un resultado más que aceptable teniendo en cuenta que la variable aquí estudiada no se puede medir.

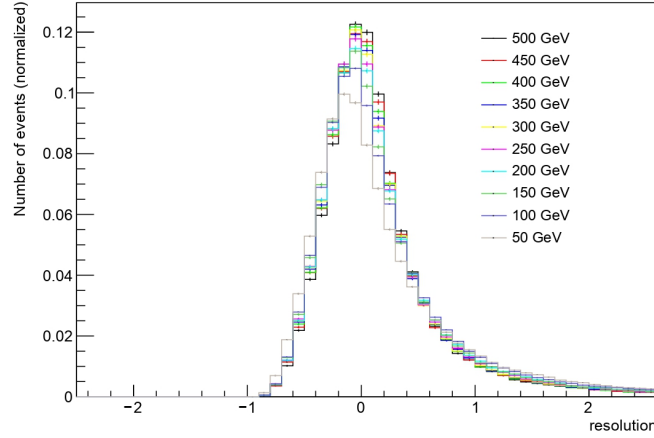
En cuanto a las distribuciones del momento del mediador ϕ , se puede ver que la red neuronal es capaz de replicar la distribución real de esta variable. No obstante, el pico de la distribución se encuentra ligeramente desplazado hacia valores de momento más altos (~ 40 GeV más alto).

Por otra parte, el histograma de la resolución también muestra una cola en la parte positiva, mientras que en la parte negativa cae a 0 completamente. Esto indica que, en los sucesos en los que la red neuronal falla a la hora de predecir, tiende a sobreestimar el momento del mediador de Materia Oscura. En el histograma 2D (especialmente en 4.2(b)) se puede apreciar una región vacía en el extremo inferior, lo que nos permite deducir que en los sucesos de bajo momento del mediador la red tiene problemas para predecir, y por norma general estima un valor superior al real.

Este fenómeno observado para bajo momento del mediador lleva a preguntarse qué sucede en función de su masa. Como se ha explicado en la sección 2.3, la información acerca del momento del ϕ se recoge principalmente en la *missing energy transverse*. Cuando la masa del mediador es muy baja, la distribución de la MET es muy similar a la del proceso $t\bar{t}$ estándar, por lo que es razonable pensar que en este caso la red neuronal tenga problemas para distinguir el efecto del mediador y por tanto funcione peor. Para comprobar esto, se ha entrenado una red diferente para cada punto de masa del ϕ , y se ha analizado su resolución (Figura 4.3).



(a) Mediador escalar.



(b) Mediador pseudoescalar.

Figura 4.3: Histogramas de resolución obtenida por la red neuronal artificial para cada punto de masa del mediador.

Aunque la diferencia no es muy notable, se puede apreciar que para masa baja del mediador el pico de resolución de la red es más bajo y más ancho, lo cual indica que efectivamente la red neuronal funciona mejor cuando la masa del mediador es alta.

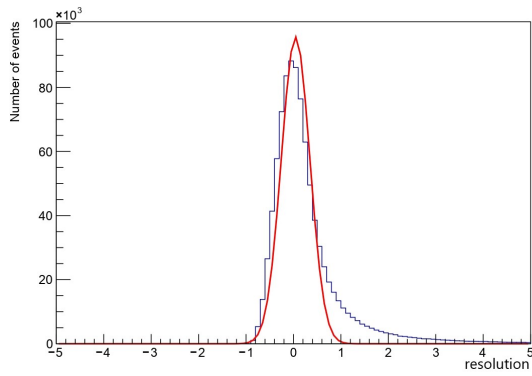
4.2. Predicción en sucesos *single top*+DM

El segundo objetivo de este trabajo es estudiar si es posible entrenar una red neuronal para predecir el momento del mediador de Materia Oscura en sucesos *single top*+DM. Este objetivo es el más relevante del estudio dado que nunca antes se ha hecho un sistema para predecir el momento del ϕ en este tipo de sucesos. En esta sección se mostrarán los resultados obtenidos y se comprobará no solo que es posible, sino que el funcionamiento de la red neuronal es muy bueno, al igual que ocurría en sucesos $t\bar{t}$ +DM.

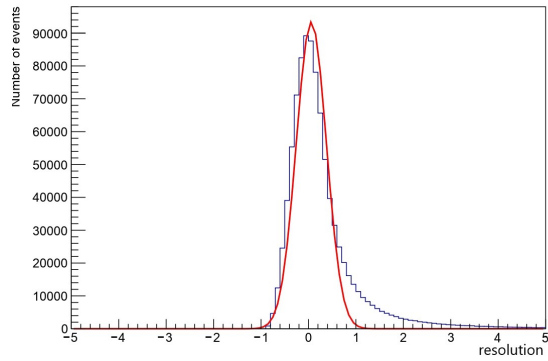
De nuevo, se ha entrenado una red neuronal para cada tipo de mediador. La arquitectura de la red ha sido la misma que antes: dos capas intermedias de N y $N/2$ neuronas, donde $N = 11$ es el número de variables (esta vez son 3 menos porque solo interviene un quark bottom). Además, la función de activación y el algoritmo de aprendizaje ha sido el mismo que anteriormente. También se ha mantenido el corte en el momento de los leptones ($p_l \geq 15$ GeV) para ajustar los sucesos a la realidad del detector.

En cuanto a las muestras disponibles para entrenar la red, la masa del ϕ varía entre 100 GeV y 500 GeV en intervalos de 50 GeV, más un punto de masa 1000 GeV. Las muestras disponían de cerca de 2,2 millones de sucesos para cada tipo de mediador, de los cuales se han empleado la mitad para entrenar la red y la otra mitad para probarla.

En la Figura 4.4 se pueden observar los resultados de la resolución obtenida por la red para ambos tipos de mediadores.



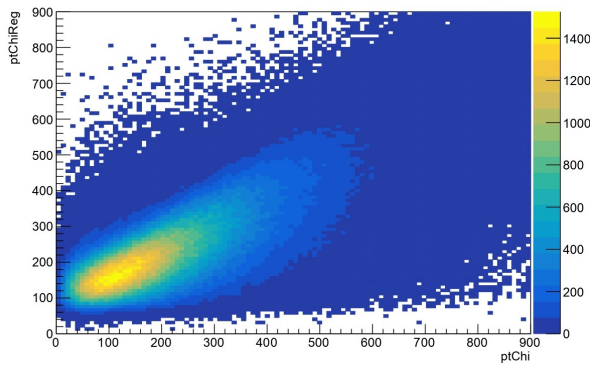
(a) Mediador escalar.
 $\mu_s = -0,1338 \pm 0,0004$ GeV
 $\sigma_s = 0,26268 \pm 0,00018$ GeV



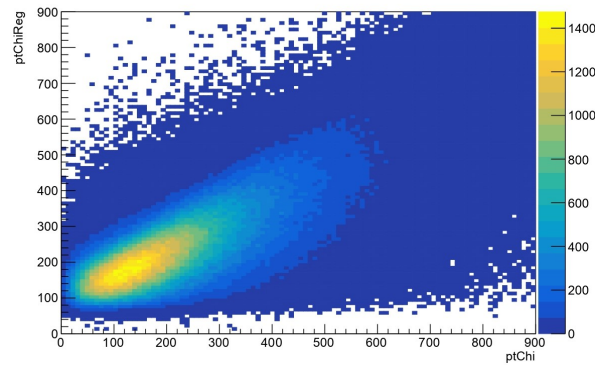
(b) Mediador pseudoscalar.
 $\mu_{ps} = -0,0715 \pm 0,0005$ GeV
 $\sigma_{ps} = 0,27092 \pm 0,00018$ GeV

Figura 4.4: Histogramas de resolución de la red neuronal artificial. Los ajustes (rojo) corresponden a dos distribuciones gaussianas, cuya media y desviación se indica arriba.

Al igual que antes, en la Figura 4.5 se muestran las distribuciones reales y reconstruidas del momento del ϕ , en forma de histograma 2D y de distribuciones superpuestas.



(a)



(b)

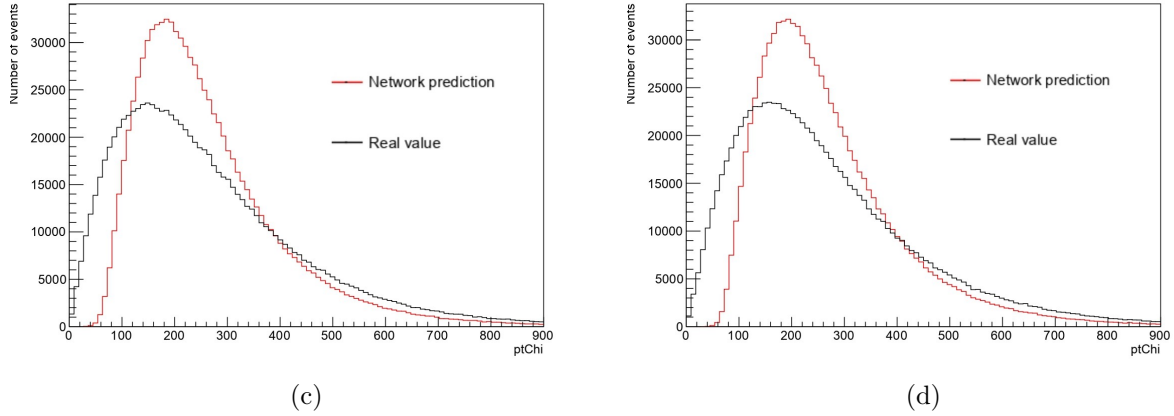


Figura 4.5: Distribución del momento del mediador de Materia Oscura. Los histogramas de la izquierda corresponden a mediador escalar y los de la derecha a mediador pseudoescalar.

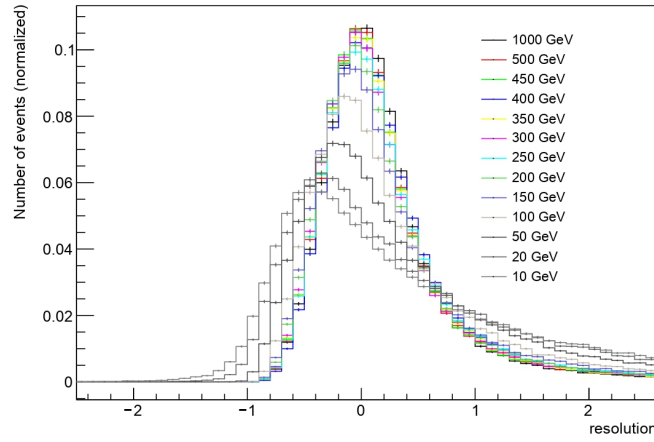
En primer lugar, en cuanto a la resolución de las ANN, la distribución se ajusta correctamente a una distribución gaussiana cuya media está en el orden de $-0,1$ GeV y su desviación es cercana a $0,3$ GeV. Si se comparan estas distribuciones con las obtenidas para sucesos $t\bar{t} + \text{DM}$, se puede ver que la desviación es muy similar pero que la media obtenida se aleja ligeramente más del 0. Esto nos indica que la red generalmente predice un momento del mediador ϕ que es un $\sim 13\%$ (mediador escalar) o un $\sim 7\%$ (mediador pseudoescalar) inferior al valor real. Estos resultados son muy aceptables teniendo en cuenta de nuevo que esta magnitud no se puede medir y que es la primera vez que se elabora un método para predecirla. Por tanto, se puede confirmar que sí es posible entrenar una red neuronal para predecir el momento del mediador de Materia Oscura en sucesos *single top* + DM.

Por otra parte, en el histograma de resolución también se observa una cola en la mitad positiva. Al igual que sucedía con sucesos $t\bar{t} + \text{DM}$, esto quiere decir que cuando la red falla al predecir el momento del ϕ tiende a predecir un momento mayor que el valor real.

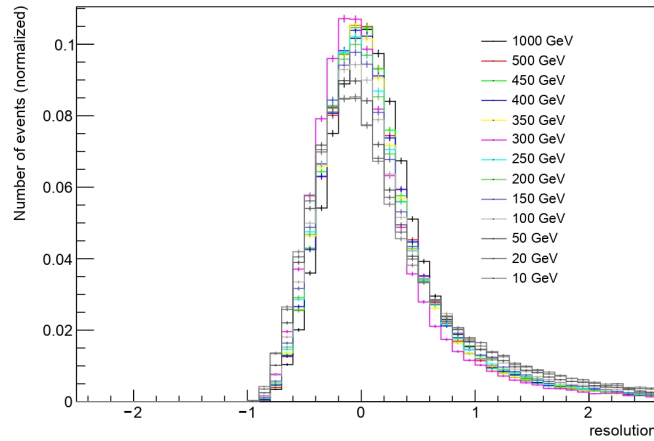
Respecto a las distribuciones del momento, hay un par de aspectos relevantes. En primer lugar, la red neuronal ha sido capaz de reproducir la forma de la distribución de momento. Sin embargo el pico de la distribución se encuentra desplazado hacia valores de momento superiores. Este desplazamiento es de aproximadamente 50 GeV. En segundo lugar, otro aspecto importante de la distribución es la región cercana a 0 GeV en el eje de momento. En esta parte, la red neuronal parece que es incapaz de predecir el momento de los sucesos. La razón de este hecho está muy relacionada con lo explicado en la sección anterior: cuando se tienen sucesos de Materia Oscura en los que el mediador lleva muy poco momento, estos sucesos son muy similares a los del Modelo Estándar, por lo que a la red le cuesta realizar una predicción buena.

Anteriormente, para sucesos $t\bar{t} + \text{DM}$ se ha concluido que la red neuronal tenía una mayor eficiencia en sucesos en los que la masa del mediador era alta, y que iba disminuyendo con cada punto de masa. De igual forma, tiene sentido preguntarse si este problema a baja masa sigue presente para sucesos *single top* + DM. En este sentido, se ha entrenado

una red diferente para cada punto de masa y se ha estudiado la resolución.



(a) Mediador escalar.



(b) Mediador pseudoescalar.

Figura 4.6: Histogramas de resolución obtenida por la red neuronal artificial para cada punto de masa del mediador.

En efecto, como se puede observar en la Figura 4.6 (especialmente en (a)), la red neuronal es muy ineficiente para los puntos de masa 10, 20 y 50 GeV. En el caso de mediador pseudoescalar la diferencia no es muy notable, pero en el caso de mediador escalar es claramente apreciable: para baja masa el pico de resolución es muy ancho y además está descentrado del 0, lo que indica que la red subestima sistemáticamente la masa del mediador.

Con todo lo anterior se puede concluir que los resultados obtenidos por las redes neuronales para predecir el momento del mediador ϕ en sucesos single top+DM son buenos. No obstante hay que destacar la limitación del método para sucesos de baja masa y bajo momento del mediador.

4.3. Comparación con el método de las elipses

Como se ha mencionado en la sección 2.3, los procesos $t\bar{t}$ estándar pueden ser reconstruidos a pesar de no disponer de las medidas de los neutrinos, ya que se dispone de una serie de ecuaciones de ligadura relacionadas con las masas de la resonancia y la medida de la MET. Así, el sistema de ecuaciones puede resolverse quedando reducido a obtener los puntos de intersección entre dos elipses en el plano formado por las componentes X e Y del momento de uno de los neutrinos [10]. En cambio, el proceso $t\bar{t} + \text{DM}$ cuenta con grados de libertad adicionales debido al momento del mediador de Materia Oscura y por lo tanto el sistema no tiene solución. Sin embargo, en el trabajo de Lantero A. [12] con el que se van a comparar los resultados, se definió una métrica para intentar estimar el momento del mediador en base a las propiedades de las elipses correspondientes. Los detalles matemáticos de este desarrollo no son el objeto de este trabajo y pueden encontrarse en las referencias dadas. No obstante, en la Figura 4.7 se muestra un esquema de este método.

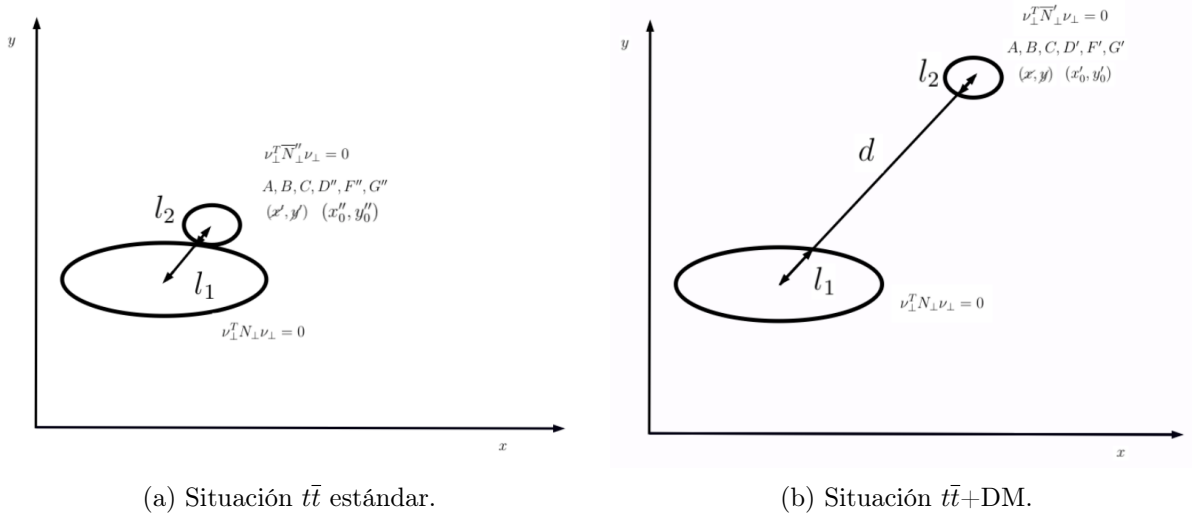


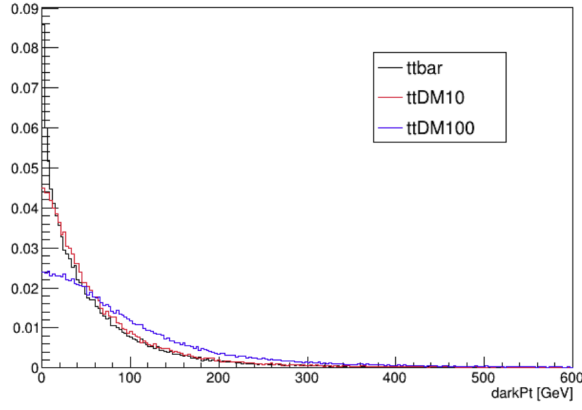
Figura 4.7: Esquema del método de obtención del dark p_T con el método de las elipses (Cortesía de Agustín Lantero [12]). En el esquema, el momento del mediador de Materia Oscura se corresponde con la magnitud d (dark p_T).

Con este método, se analizaron numerosos sucesos obtenidos a partir de una simulación de Monte Carlo para tratar de obtener el valor del dark p_T . Los resultados obtenidos fueron unas distribuciones de dark p_T diferentes a la real: para los sucesos de bajo momento, la distribución real cae a 0, mientras que la obtenida con este método crece.

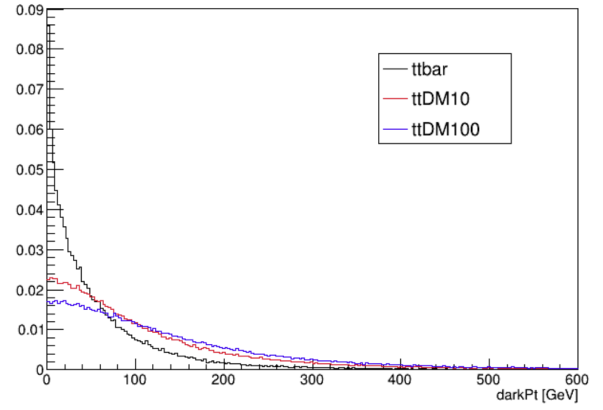
Para lograr que la distribución del momento de las elipses sea similar a la real, hay que introducir de forma artificial un parámetro llamado factor de solapamiento, R , que es un indicador del tamaño y la distancia de separación de las elipses de momento, y permite discriminar los sucesos $t\bar{t} + \text{DM}$.

$$R = \frac{l_1 + l_2}{d}$$

Valores pequeños de R permiten discriminar sucesos $t\bar{t} + \text{DM}$, ya que cuanto menor sean las elipses y más separadas estén, menor será la incertidumbre en el momento de ϕ .



(a) Todos los eventos



(b) Todos los eventos

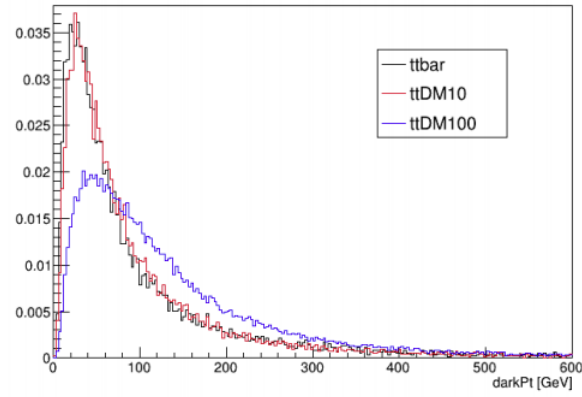
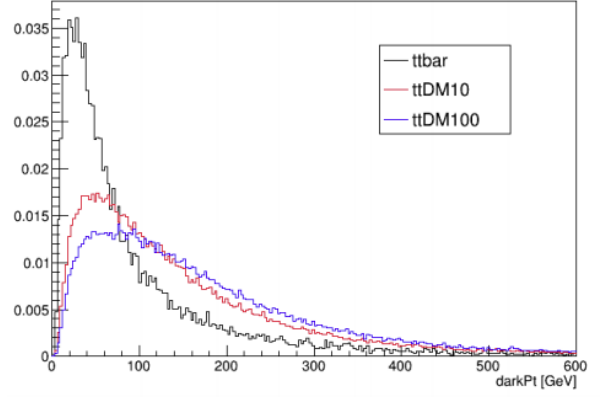
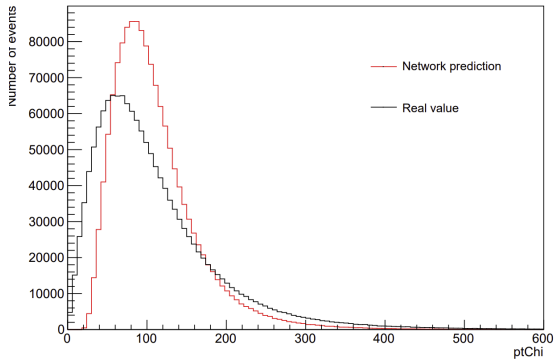
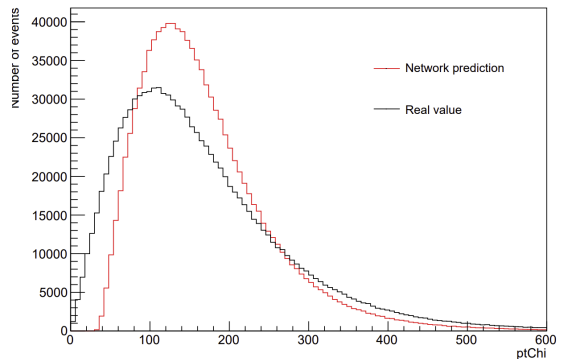
(c) Eventos $R < 0,9$ (d) Eventos $R < 0,9$

Figura 4.8: Distribución del momento del mediador de Materia Oscura. obtenido con el método de las elipses (Cortesía de Agustín Lantero [12]). Los dos histogramas de la izquierda corresponden a mediador escalar y los de la derecha a pseudoescalar.



(a) Mediador escalar. 100 GeV.



(b) Mediador pseudoescalar. 100 GeV

Figura 4.9: Distribución del momento del mediador de Materia Oscura. obtenido con las redes neuronales. Los histogramas corresponden solo a sucesos con masa del mediador 100 GeV.

A diferencia de estos resultados, con el método de las redes neuronales artificiales se ha podido replicar sin problema la distribución real del momento de ϕ , sin tener que recurrir al factor de solapamiento. Este hecho garantiza que de alguna forma la información del factor de solapamiento está contenida de forma implícita en la red.

Comparando la Figura 4.9 con la Figura 4.8, se ve que para una masa de mediador de 100 GeV la red neuronal es mucho más eficaz a la hora de predecir el momento que el método analítico de las elipses.

Además, el factor de solapamiento introduce el problema de que se desecha parte de la estadística de datos de la que se dispone, ya que se seleccionan solo aquellos sucesos para los que la reconstrucción es buena. Sin embargo, con la red neuronal se consigue una buena reconstrucción para todos los sucesos.

Capítulo 5

Conclusiones

En este trabajo se ha explorado la capacidad de las redes neuronales artificiales para predecir el momento de una partícula mediadora de Materia Oscura. En concreto, se han estudiado dos tipos de topologías diferentes: $t\bar{t}$ +DM y *single top*+DM. Este problema se enmarca en el contexto de una búsqueda activa de Materia Oscura que se lleva a cabo en el experimento CMS.

Para concluir este trabajo, es conveniente volver sobre los objetivos presentados en la sección 2.5, y analizar si se han cumplido o no.

En primer lugar, se ha comprobado que sí es posible utilizar una red neuronal para predecir el momento del mediador de Materia Oscura en sucesos $t\bar{t}$ +DM. Además, los resultados obtenidos se han comparado con el método analítico de las elipses [12], y se ha observado que el poder de predicción de la red neuronal es superior al método analítico.

Por otra parte, también se ha comprobado que sí es posible entrenar una red neuronal para predecir el momento del mediador ϕ en sucesos *single top*+DM. Con ello, el sistema aquí presentado es el primero capaz de reconstruir el momento del ϕ en este tipo de sucesos. Esto tiene una gran importancia porque en el futuro puede ser aplicado sobre sucesos reales para discriminar el proceso *single top*+DM y caracterizarlo.

No obstante, este método basado en redes neuronales tiene limitaciones, las cuales no hay que pasar por alto. La más importante de ellas es la limitación observada en sucesos de baja masa y/o bajo momento del mediador. Aquí, de forma general, las redes neuronales han predicho un valor de momento superior al real, lo cual tiene dos efectos sobre los resultados presentados: en el histograma de resolución aparece una cola en la mitad positiva y en la distribución de momento no se predicen valores inferiores a un cierto límite.

Una posible interpretación de esto está directamente relacionada con las elipses de momento de los neutrinos mencionadas anteriormente. Cuando el momento del mediador es pequeño, es posible que su valor esté dentro del rango de incertidumbre que conllevan dichas elipses. En este caso, el efecto del mediador sobre la MET es pequeño y la red neuronal no es capaz de predecir el momento del ϕ de forma satisfactoria.

Otra discrepancia observada, especialmente en sucesos *single top*+DM, es que para baja masa del mediador (< 100 GeV), la resolución de la ANN es mejor para mediador pseudoescalar que para mediador escalar. Esto puede estar relacionado con la propia estructura angular del mediador. Los mediadores escalar y pseudoescalar tienen distinta estructura angular debido a la introducción de la matriz de Dirac γ^5 en el vértice de la interacción. En particular, se ha observado que la interacción escalar produce una cantidad menor de MET que la pseudoescalar debido al antialineamiento de los neutrinos con la partícula mediadora de Materia Oscura, produciendo una cierta cancelación. Entender los detalles de este fenómeno es algo que queda como trabajo futuro. En cualquier caso, este aumento de la MET en el caso pseudoescalar hace que la MET se diferencie más de la del proceso $t\bar{t}$ estándar, y esto proporciona mejor información a la red para realizar el aprendizaje.

Con todo lo anterior se puede concluir que las redes neuronales son una excelente herramienta para la predicción del momento del mediador ϕ , y que es una línea interesante para explorar y aplicar en la búsqueda de materia de oscura de CMS. Más adelante se podrá explorar esta técnica sobre sucesos reales obtenidos a partir de colisiones protón-protón en el LHC, y con ello tratar de discriminar y caracterizar sucesos de Materia Oscura.

Bibliografía

- [1] L. Kelvin, *Baltimore lectures on molecular dynamics and the wave theory of light*. London : C. J. Clay and sons; Baltimore, Publication agency of the Johns Hopkins university, 1904.
- [2] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helvetica Physica Acta*, vol. 6, pp. 110–127, 1933.
- [3] K. Freese, “Status of Dark Matter in the Universe.” <https://ned.ipac.caltech.edu/level5/Sept17/Freese/Freese3.html>.
- [4] “Modelo Estándar de la Física de Partículas.” https://es.wikipedia.org/wiki/Modelo_est%C3%A1ndar_de_la_f%C3%ADsica_de_part%C3%ADculas#:~:text=El%20modelo%20est%C3%A1ndar%20de%20la,como%20entes%20irreducibles%20cuya%20cinem%C3%A1tica.
- [5] M. Thomson, *Modern Particle Physics*. New York: Cambridge University Press, 2013.
- [6] “The Large Hadron Collider.” <https://home.cern/science/accelerators/large-hadron-collider>.
- [7] “CMS.” <https://home.cern/science/experiments/cms>.
- [8] “CMS Detector.” <https://cms.cern/detector>.
- [9] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *Journal of High Energy Physics*, vol. 2014, Jul 2014.
- [10] B. A. Betchart, R. Demina, and A. Harel, “Analytic solutions for neutrino momenta in decay of top quarks,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 736, p. 169–178, Feb 2014.
- [11] L. Sonnenschein, “Analytical solution of $t\bar{t}$ dilepton equations,” *Phys. Rev. D*, vol. 78, Oct 2008.
- [12] A. Lantero, “Mejora de la discriminación de señal y fondo en una búsqueda de materia oscura producida en asociación con un par de quarks top-antitop,” Sep 2018.

- [13] “Prediction of wind pressure coefficients on building surfaces using Artificial Neural Networks - Scientific Figure on ResearchGate.” https://www.researchgate.net/figure/Artificial-neural-network-architecture-ANN-i-h-1-h-2-h-n-o_fig1_321259051.
- [14] G. Bertone and D. Hooper, “History of dark matter,” *Reviews of Modern Physics*, vol. 90, Oct 2018.
- [15] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Bakes, T. Carli, O. Cohen, A. Christov, D. Dannheim, K. Danielowski, S. Henrot-Versille, M. Jachowski, K. Kraszewski, A. K. Jr., M. Kruk, Y. Mahalalel, R. Ospanov, X. Prudent, A. Robert, D. Schouten, F. Tegenfeldt, A. Voigt, K. Voss, M. Wolter, and A. Zemla, “TMVA - Toolkit for Multivariate Data Analysis,” 2007.
- [16] P. Zyla *et al.*, “Review of Particle Physics,” *PTEP*, vol. 2020, no. 8, p. 083C01, 2020.